{PRESCRIBE FORMAT FOR SUBMISSION OF PROPOSAL FOR TOPIC APPROVAL IN THE FACULTY OF SCIENCE AND TECHNOLOGY}

PROPOSAL FOR THE TOPIC APPROVAL TO BE SUBMITTED TO THE UNIVERSITY OF MUMBAI FOR THE PHD DEGREE IN THE SUBJECT OF SCIENCE & TECHNOLOGY

Tittle of the Topic	Design and Development of Deep Learning Model for Brain Tumour Classification Using MRI Images
Name of the student	Ravindra Sugdeo Sonavane
Basic qualification	M.E. & B.E. in E.X.T.C.
In the case of Ph.D. student year of passing of PET/NET/SET/M.PHIL/GPAT/GATE. (Teaching Experience)	Gate Qualified in March' 2025 with Score 24.67 Exp: 10 years & 6 Months
Name of the Research /Supervisor /Guide	Dr. Mansing Rathod
Date of the Superannuation /Retirement of the Research Supervisor guide	6/6/2034
Name of the research Centre	K.J. Somaiya Institute of Technology, Sion
Signature of the student	Ravi S
Signature of the Research Supervisor /Guide	MRathod 16/2/26

Date: 16/2/26

Somaiya Institute of Technology Sion Mumbai 400 022

Seal of College/ Department

Signature of
Head/Director/Principal
Research Centre

CERTIFICATE OF APPROVAL

This is to certify that **Ravindra S. Sonavane** has delivered a seminar on **"Design and Development of Deep Learning Model for Brain Tumour Classification Using MRI Images"** on 16th February 2026 and submitted a report in **K J Somaiya Institute of Technology**, Sion, Mumbai, in the partial fulfilment of the degree of PhD in Science and Technology (Information Technology).

Rathod
(Dr. Mansing Rathod)
Supervisor
(Name and Signature)

Subject Expert Internal
(Name and Signature)

Subject Expert External
(Name and Signature)
Dr. A. M. Kanthe

Head of the Department
(Name and Signature)

PhD Coordinator
(Name and Signature)

Principal

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity, and have not misrepresented, fabricated, or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been appropriately cited or from whom proper permission has not been taken when needed.

Mr. Ravindra Sugdeo Sonavane

Name of the Student

Admission No:2220250739

Date: 16/2/26

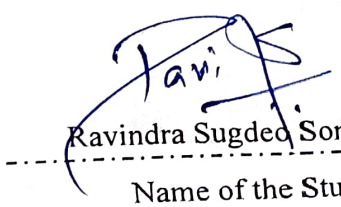
Place: Mumbai

ACKNOWLEDGEMENT

It gives a great pleasure to acknowledge our deep sense of gratitude to present my Special Topic Report project titled: "Design And Development of Deep Learning Model for Brain Tumour Classification Using MRI Images". I would like to give sincere thanks to our Principal Dr. Vivek Sunnapwar, my HOD Dr. Radhika Kotecha and my Guide Dr. Mansing Rathod for giving me the opportunity to present this topic. His dynamism, vision, and sincerity have deeply inspired us. He has guided to present this work as clearly as possible.

I am thankful to our PhD Co-ordinator Prof. Hari Ram Chavan for this whole hearted support and affectionate encouragement.

I am also thankful to my institution for offering the necessary resources and learning environment. The encouragement from my peers played a vital role in motivating me. I appreciate the reference materials and online resources that enhanced my understanding and conceptual knowledge. This acknowledgement is a small reflection of my deep appreciation for everyone involved.



Ravindra Sugdeo Sonavane

Name of the Student

Admission No:2220250739

TABLE OF CONTENTS

Sr. No.	Content	Page No.
1	Title	1
2	Aim of Study	1
3	Objectives	1
4	Study Background (Literature Review) and Significance of the Study	1
5	National Status	9
6	International Status	10
7	Study Hypotheses	11
8	Study Rationale	12
9	Research Methodology	12
10	Statistical Plan	15
11	Limitations of the Study	17
12	Expected Outcomes of the Study	17
13	Conclusion	17
14	Bibliography	18

LIST OF FIGURES

Sr. No.	Content	Page No.
1	Block diagram of the proposed system	13

LIST OF TABLES

Sr. No.	Content	Page No.
1	Summary of survey on brain MRI detection and classification Using Deep Learning Algorithms	7
2	Statistical Plan with Duration with Activity	15

1. **TITLE:** Design And Development of Deep Learning Model for the Brain Tumour Classification Using MRI Images

2. AIMS OF THE STUDY

In order to create an end-to-end deep learning model in imaging to segment brain tumors and classify them into multiple classes using MRI images, to enhance the accuracy and interpretability of the results, the Attention U-Net that is an attention-based CNN, evolutionary optimization (GWO/AGWO), and Explainable AI should be incorporated.

3. OBJECTIVES OF THE STUDY

1) To conduct a systematic literature survey on brain tumor segmentation and multi-class classification using MRI images.
2) To design and develop an optimized attention- based UNET deep learning architecture for brain tumor segmentation.
3) To design and develop the attention-based CNN algorithm with optimization methods for multi-class Brain MRI image classification
4) To design and develop a hybrid Deep Learning Model for brain tumor classification using MRI images, and validate the results

4. STUDY BACKGROUND (LITERATURE REVIEW) AND SIGNIFICANCE OF THE STUDY

4.1. Literature Review

Abraham et al. [1] (2026) provide a systematic review of explainable artificial intelligence (XAI) for brain tumor detection and classification. The methodology is based on a PRISMA-based review program, which includes CNNs, transformers, hybrid architectures, and XAI methods, including Grad-CAM, LIME, and SHAP. According to the research findings, DenseNet121 and ViT-based models achieved over 97% accuracy and were more interpretable. However, the review creates a gap concerning computational overhead, as well as a lack of narrow real-time deployment and standardized measures of explainability as implemented in clinical practice.

The proposed automated framework by Khan et al. [2] (2025) to classify brain tumors based on MRI images aims to minimise diagnostic errors and unnecessary radiation exposure. The authors proposed a hybrid deep learning system combining CE-EEN-B0 and ResGANet with ensemble-based feature selection. The suggested method was compared with the models such as InceptionV3, Vision Transformer, DenseNet, and ResNet with an 99.11% accuracy with a higher precision, recall and

Cohen Kappa score. Although this is high performance, it was tested using small datasets and also problems associated with multi-center generalization, interpretability, and clinical deployment in real time were not mentioned.

The presented work (Immaculate Joy S. et al. [3]) (2025) has introduced an in-depth CNN-based multi-grade brain tumor classification system focusing on extensive data augmentation in the attempt to address the problem of an imbalanced dataset. The techniqueology used several convolutional layers of batch normalization and violent data augmentation methods which dramatically increased the data set. The model attained a precision of 99 percent and validation accuracy of 91.5 percent. Nevertheless, cross-dataset validation and resistant to scanner variability was not studied, which restricted clinical generalizability of the study.

The systematic review by Fahim et al. [4] (2025) is based on PRISMA and includes a review of the latest use of deep learning in brain tumor detection, classification, and segmentation of MRI images. The review classified methods as CNNs, transformers and hybrid models, and datasets, preprocessing methods and evaluation measures. Although the research reported the effectiveness of hybrid architectures, it also reported major gaps such as lack of interpretability, inadequate clinical validation and inadequate robustness in varied imaging setting.

Kar and Singh [5] (2025) suggested MBTC-Net, which is a multimodal brain tumor classification framework that combines the MRI and CT images. The extraction of features was done by EfficientNetV2B0, and multi-head attention was employed to ensure the extraction of contextual dependencies. The model had an accuracy of up to 99.34 in various classification tasks. However, the method is based on sparse publicly available multimodal data and fails to consider uncertainty estimation or analysis in terms of deployment.

Aiya et al. [6] (2025) suggest a better hybrid deep learning network that would be composed of VGG16, attention, and Grad-CAM visualization on brain tumor. The given methodology is premised on applying transfer learning, optimized hyperparameters, and end-to-end learning to differentiate between MRI images of glioma, meningioma, pituitary, and non-tumor tumors. The model was 99 percent accurate in classifying and is more clinical explainable. Nevertheless, it is also computationally intensive in the architecture as well as has never been tested on large multi-centered datasets and thus it is small scale limited.

Xiong [7] (2025) introduces a ressauxu-net, an adapted U-Net that is used to segment and classify multimodal brain tumors. The approach combines the ResNet-based down-sampling, squeeze-

excitation networks as recalibration, and hybrid Dice-cross-entropy loss to meet the issues of class-imbalance. The similarity coefficient between the proposed model and Dice on the BraTS2018 and BraTS2019 data was more than 0.96. The strategy is however, mostly segmentation-oriented and does not consider classification-only pipelines or real-time deployment constraints.

In a different research, Wong et al. [8] (2025) provide an automated structure of brain tumor sorting by the deep learning method using MRI images, which has various classifications. The methodology is the pre-trained VGG16 convolutional neural network with fine-tuning and extensive data augmentation that was trained on a mixed dataset of 17,136 MRI scans including glioma, meningioma, pituitary tumors and normal cases. The classification accuracy is experimental and is 99.24, which is also better than certain of the available CNN-based models. However, the technique does not provide tumor segmentation, attention and robustness analysis in heterogeneous clinical scanners and, therefore, limits the level of practical application.

Ali et al. [9] (2025) present the hybrid framework of multi-class brain tumor MRI segmentation and classification based on deep learning and machine learning. This methodology includes edge-refined binary histogram segmentation, hybrid feature extraction, feature optimization using correlation, and evaluation using multiple classifiers using ten fold cross-validation. The best was 98.61 when it was done by the random committee classifier. Nevertheless, it is less scalable and incapable of end-to-end learning because of such features as handcrafted and multi-stage pipeline.

The article by Aiya et al. [10] (2025) provides the improved hybrid deep learning model incorporating attention-based mechanisms and clinical explainability during the brain tumor classification operation. It is a backbone of VGG16 that has a custom attention layer and Grad-CAM as a visual interpretability technique. There is a lot of transparency in the proposed model and it is 99 percent accurate in classification. However, the model is computationally intensive and has not been used on large multi-centre clinical data.

The article by Celik Ertugrul et al. [11] (2025) presents a hybrid brain tumor detection model that detects it using MRI and X-ray. The methodology compares various deep learning models that include CNN, ResNet-50, DenseNet-121, AlexNet and XGBoost and finally performs majority voting based fusion strategy. The ensemble technique achieved 100 percent post fusion accuracy. Though it is high performance the study says nothing about model interpretability, clinical feasibility and resistance to noisy medical images.

An attention-guided convolutional neural network of edge-computing based brain tumor classification, as introduced by Babar et al. [12] (2025), is shallow. The methodology also presents an attention based lightweight CNN architecture with model compression and Grad-CAM visualization. The suggested model demonstrated average accuracies of more than 95% on various datasets with a large decrease in the inference latency. Nevertheless, the segmentation of tumors and its performance when there is a severely skewed class distribution were not examined.

Arif et al. [13] (2025) introduce an automated detection system of brain tumors, which is based on ensemble classification and localization with YOLOv3. The approach combines handcrafted GLCM properties with SVM and KNN classifiers with the use of YOLOv3 to detect tumors in a bounding-box. YOLOv3 model had an accuracy of 97.80 and the mean IoU of 0.65. The method however does not provide accurate tumor segmentation and uses manually engineered features, hence restricting its flexibility and strength.

The article by Zine-dine et al. [14] (2025) is the detailed overview of the machine learning and deep learning approaches to brain tumor segmentation and classification. The methodology surveys include analysis of feature-extraction methods, such as GLCM, LBP, DWT, PCA, and CNN based models as well as gives a performance comparison of these methods. The review finds that deep learning methods are always better when compared to classical methods. However, the issues of computational complexity, imbalance of data, and non-interpretability are still unsolved.

Ghosh et al. [15] (2025) provide an optimized deep learning system to detect and classify the brain tumor through the MRI images. The methodology will involve image preprocessing, image segmentation, and analysis of various deep learning structures to select the best one. The streamlined model resulted into better reliability of diagnostic and classification. Nonetheless, there was no incorporation of attention mechanisms and multimodal imaging data and that restricted the richness of diagnostic aspects.

Mjahad and Rosado-Muñoz [16] (2025) introduce an aberration based brain tumor detection algorithm which makes use of CNN based feature representation along with One-Class SVM. The methodology deals with the issue of class imbalance by training mostly on normal MRI images with the use of DenseNet121, VGG16, and MobileNetV2 feature extractors. CNN-only model had an accuracy of 97.83. The method is however not explicit in tumor segmentation and it does not have the explainability mechanisms necessary to trust it clinically.

The article by Gomez-Guzmana et al. [17] (2025) is a submission of a more advanced framework of brain tumor multi-classification based on pre-trained convolutional neural networks and transformer-based architectures on MRI scans. The methodology compares DeiT3, Xception, Inception-v4, and Swin-Transformer models, which are transferred and have higher preprocessing and stratified five-fold cross-validation in the Msoud Brain Tumor MRI dataset. According to experimental findings, the Swin-Tiny transformer had reached an ultimate accuracy of 99.24 per cent whilst being robust with smaller training information and realistic-time performance on edge-based devices. Nonetheless, the research relies more on classification and fails to incorporate fine-grained tumor segmentation or clinical explainability mechanisms to improve the transparency of decisions.

Agarwal et al. [18] (2024) designed a hybrid automated framework to detect and categorize brain tumors with reference to contrast-enhanced MRI images. The method involved an auto-contrast optimization element with the transfer learning based on the Inception-V3 model. The proposed system was found to have the highest accuracy of 98.89 surpassing some of the baseline CNN architectures. Nevertheless, the use of handcrafted contrast improvement and the lack of robustness test in the conditions of actual clinical noise is still the main limitations.

The research of Anantharajan and colleagues [19] (2024) created a hybrid of deep learning and machine learning that incorporates adaptive contrast enhancement, fuzzy C-means segmentation, GLCM feature extraction and ensemble deep neural SVM classifier. Its specificity and accuracy of the procedure was 97.93. However, the multi-stage pipeline is a conceptually involved and not particularly PC-friendly algorithm and systems that relies largely on manually crafted features which limits its scaling and overall optimization.

De Benedictis et al. [20] (2024) proposed a machine learning of topological data analyses is a explainable model of brain tumor classification. The methodology involved low-rank Tucker decomposition, persistent homology, and ensemble classifiers in order to increase interpretability. The precision of the system was 97.28. The method was not however compared to the new end-to-end deep learning systems and real-time practicability was not analyzed.

Reddy et al. [21] (2024) proposed a CNN driven brain tumor detection and classification model that uses MRI images, and compared it with ResNet-50, VGG-16 and Inception-V3. The proposed model achieved 93.3 percent accuracy and AUC of 98.43. However, the performance was poor in comparison with the presence of new hybrid and attention-based architectures and segmentation-based classification was not considered.

Brain tumor detection and classification framework developed by Mithun and Jawhar [22] (2024) is based on a YOLO-NAS architecture and a segmentation model based on U-Net. The approach employed hybrid anisotropic diffusion filtering, En-DeNet segmentation, and YOLO-NAS in the detection of multi-classes of tumors. The model was accurate in 99.7 percent with high sensitivity and specificity. Despite the fact that the performance is excellent, the framework is computationally expensive and fails to interpret and model the uncertainty, which are required to win clinical confidence.

In Ahmed et al. [23] (2024), the hybrid Vision Transformer-Gated Recurrent Unit (ViT-GRU) model is offered to identify and classify brain tumors based on the MRI data in Southern Bangladesh. ViT is used together with GRU to learn the spatial features in a sequence, and XAI techniques, including attention maps, SHAP and LIME are used in the methodology. The given model was the most interpretable and less prone to class imbalance with a maximum accuracy of 98.97. The framework has however increased complexity of computation and has not been proved in other regions except specific clinical datasets.

The authors developed a parallel deep convolutional neural network (PDCNN) architecture (2023) [24] that was designed to extract global and local features of MRI images. There were also two CNNs that were parallel to each other with different kernel sizes, dropout, and batch normalization. On a number of datasets, the model was found to classify over 97 percent. Despite having a lot of promise, the architecture lacks the attention mechanism and was not tested on highly unbalanced real-world clinical data.

Neamah et al. [25] (2023) is a systematic review of the deep learning based techniques of brain tumor detection and classification based on MRI images. The comparison of architectures proposed in 2019-2022 on CNNs, datasets, evaluation, and architecture trends is the approach in the methodology. It is stated in the review that deep learning models are better than the old machine learning techniques. It is however discovered to have significant flaws, particularly inability to be explained, overfitting on a small sample size, and failure to be clinically validated.

Babu Vimala et al. [26] (2023) study provides a hybrid deep learning brain tumor classification model on the basis of transfer learning on the EfficientNet frameworks. The approach optimizes EfficientNet-B0 to B4s on the CE-MRI Figshare data, through data enhancement and Grad-CAM to enhance the visual sensitivity. EfficientNet-B2 model achieved the highest level of accuracy 99.06 percent and a high degree of precision and recall. The method, though doing well, requires a lot of computing memory and does not have tumor segmentation and cross-dataset transfer.

Table 1: Summary of survey on brain MRI detection and classification Using Deep Learning Algorithms

Ref	Dataset Used	Key Steps (Preproc / Segmentation / Features)	Classifier / Model	Dataset / Performance	Research Gap
[1]	Multiple public MRI datasets (reviewed)	PRISMA-based review; CNNs, ViTs, hybrid DL; XAI (Grad-CAM, LIME, SHAP)	CNNs, ViTs, Hybrid DL + XAI	Review; Acc >97% reported	High computational cost; no real-time focus; no standardized XAI metrics
[2]	Limited public MRI datasets	Preprocessing; ensemble feature selection	CE-EEN-B0 + ResGANet	Acc = 99.11%	No multi-center validation; interpretability and real-time use not addressed
[3]	Augmented MRI dataset	Aggressive augmentation; BN	Deep CNN	99% (train), 91.5% (val)	No cross-dataset validation; scanner variability ignored
[4]	Multiple MRI datasets (reviewed)	PRISMA review; methods, metrics, preprocessing	CNNs, Transformers, Hybrid DL	Review	Lack of interpretability; weak clinical validation
[5]	MRI + CT (limited multimodal sets)	Feature extraction; attention fusion	EfficientNetV2B0 + MHA	Acc up to 99.34%	Limited multimodal data; no uncertainty modeling
[6]	Kaggle MRI dataset	Transfer learning; attention; Grad-CAM	VGG16 + Attention	Acc ≈ 99%	Computationally intensive; no multi-center testing
[7]	BraTS 2018, BraTS 2019	ResNet down-sampling; SE blocks; hybrid loss	ResSAXU-Net	Dice > 0.96	Segmentation-focused; real-time/classification-only not evaluated
[8]	Combined MRI dataset (17,136 images)	Augmentation; fine-tuning	VGG16 CNN	Acc = 99.24%	No segmentation/attention; scanner robustness untested
[9]	Public MRI datasets	ER-BHS segmentation; handcrafted features	ML (Random Committee)	Acc = 98.61%	Multi-stage; not end-to-end
[10]	Kaggle MRI dataset	Attention-based features; Grad-CAM	VGG16 + Custom Attention	Acc ≈ 99%	High compute; no extensive multi-center validation
[11]	MRI + X-ray datasets	CNN features; ensemble voting	CNN, ResNet50, DenseNet121, XGBoost	Acc = 100% (fusion)	No explainability; clinical feasibility unclear
[12]	Multiple MRI datasets	Lightweight preprocessing; attention; compression	Shallow Attention CNN (Edge)	Acc > 95%	No segmentation; imbalance not analyzed
[13]	Public MRI datasets	GLCM features; YOLO localization	YOLOv3 + SVM/KNN	Acc = 97.8%; IoU = 0.65	Handcrafted features; coarse localization
[14]	Multiple MRI datasets (reviewed)	Survey of ML/DL features	ML, CNN, ViT (Review)	Review	Complexity; imbalance; limited XAI
[15]	Public MRI datasets	Preprocessing; segmentation; model comparison	Multiple DL models	Improved accuracy	No attention; no multimodal data
[16]	Public MRI datasets	CNN features; anomaly detection	CNN + One-Class SVM	Acc = 97.83%	No segmentation; lacks XAI
[17]	Msoud Brain Tumor MRI dataset	Transfer learning; ViTs; preprocessing	Swin Transformer, CNNs	Acc = 99.24%	No segmentation; XAI absent

[18]	Contrast-Enhanced MRI	Auto-contrast enhancement; FL	InceptionV3	Acc = 98.89%	Handcrafted preprocessing; noise robustness untested
[19]	Public MRI datasets	Contrast enhancement; FCM; GLCM	DNN + SVM Ensemble	Acc = 97.93%	Computationally heavy; not end-to-end
[20]	Public MRI datasets	Topological features; tensor decomposition	ML Ensemble	Acc = 97.28%	Not compared with modern DL; no real-time analysis
[21]	Public MRI datasets	CNN feature extraction	Custom CNN vs ResNet/VGG	Acc = 93.3%; AUC = 98.43%	Lower performance; no segmentation
[22]	Public MRI datasets	Anisotropic filtering; U-Net segmentation	YOLO-NAS + U-Net	Acc = 99.7%	Computationally heavy; no XAI
[23]	Regional MRI dataset (Bangladesh)	ViT features; GRU; XAI	ViT-GRU + SHAP/LIME	Acc = 98.97%	High complexity; region-specific validation
[24]	Multiple MRI datasets	Parallel conv paths; BN; dropout	PDCNN	Acc > 97%	No attention; imbalance not tested
[25]	Multiple MRI datasets (reviewed)	Systematic DL review	CNNs (Review)	Review	Overfitting; limited XAI
[26]	CE-MRI Figshare	Transfer learning; augmentation; Grad-CAM	EfficientNet (B0–B4)	Acc = 99.06%	High compute; no segmentation; weak generalization

The above literature review presents the following research gap:

The updated review of latest literature on the brain tumor detection, classification, and segmentation based on MRI shows that there is a tremendous advancement in the classification accuracy as most deep learning and hybrid-based models reach over 97 percent of accuracy. CNN-based architectures, transformer architectures, attention mechanisms, and hybrid architectures have been shown to be quite powerful in discriminating amongst various tumor types. Nevertheless, in spite of these innovations, there are still a number of important and endemic research gaps.

- Existing segmentation models such as U-Net variants achieve high Dice scores but lack robust multi-scale attention refinement and cross-dataset generalization [4], [7], [25].
- Many studies address segmentation and classification as separate tasks; unified end-to-end frameworks integrating both remain limited [4], [8], [17].
- CNN-based classification models report high accuracy (>97%), yet often rely on standard backbone architectures without task-specific optimization [6], [26].
- Overfitting and limited validation across heterogeneous, multi-center MRI datasets persist in many classification studies [3], [21], [25].
- Hybrid deep learning frameworks frequently depend on manual hyperparameter tuning rather than automated evolutionary optimization strategies [2], [9], [19].

- Computational complexity and lack of convergence stability analysis are reported in several hybrid and transformer-based models [22], [23].
- Explainable AI techniques (Grad-CAM, SHAP, LIME) are predominantly applied qualitatively, with limited quantitative interpretability assessment [1], [4], [25].
- Uncertainty estimation and confidence-aware decision modeling remain largely unexplored in brain tumor diagnostic systems [5], [22].

4.2.Significance of the Study

Diagnosis of brain tumor is a clinical problem of critical and time sensitive nature in which early and precise diagnosis has a direct effect on the planning of treatment and survival of the patient. Although the current state of deep-learning-based medical image analysis has made significant progress, current methods are associated with disjointed pipelines, narrow generalization, a lack of explanations, and excessive complexity, which make them hard to implement in real-life clinical settings. In this regard, this current study has a great academic, clinical, and technological value.

The proposed research offers a contribution to clinical implications through the development of an end-to-end solution that combines brain tumor segmentation and multi-classification of MRI images. This combined method increases the coherence of the diagnostic process as it is used to localize the tumor regions and determine the type of tumor at the same time providing the radiologists with more accurate and consistent information. Explainable AI mechanisms also enhance clinical trust because clinicians can interpret and check the predictions of the model and make it less black-box, as provided by deep learning systems.

Technologically, the study contributes to technology as it overcomes such critical constraints of current models as the poor generalization within heterogeneous datasets and excessive computational costs. By confirming the suggested framework to work on multi-source MRI imaging data and penalizing it to achieve computational efficiency, the study will be able to implement it in real-time using scarce resources in a clinical setting, such as resource-minimal hospitals and edge-computing devices.

Academically and research-wise, the work contributes to the literature in a systematic manner as it fills the gaps that have been discovered in recent literature on end-to-end learning, standardized explainability and robustness evaluation. The suggested methodology and experimental results can be used as a reference framework in further studies and investigations in the medical image analysis, explainable AI, and clinical decision support systems.

The research has a crucial role to play in the development of trustworthy, interpretable and implementable artificial intelligence solutions in brain tumor diagnosis, and has the potential of

enhancing accuracy in diagnosis, lessening the load of clinicians, and aiding the timely and informed clinical decision-making.

5. NATIONAL STATUS

Diagnosing and categorizing brain tumors with the assistance of medical imaging became a burning issue in India because researchers have achieved great advances in the domain of deep learning, additional information about MRI is accessible and more medical workers are interested in computer-aided diagnostics (CAD) systems. Some research on the implementation of transfer learning models based on convolutional neural networks (CNNs) and machine learning hybrid algorithms have been done on the national scale in the context of automated brain tumor detection in MRI images.

Indian scientists have offered a vast variety of methods such as CNN-based classification systems, U-Net and its variations to tumor segmentation, and composite frameworks of handcrafted feature extraction with deep learning classifiers. BraTS, Figshare CE-MRI and Kaggle brain tumor datasets are publicly available datasets which are widely used in national studies to conduct training and validation. The classification accuracies of many of these methods claim to be very high, above 95 percent, showing that deep learning may be used to diagnose brain tumors.

Nevertheless, there are also some weaknesses of the national research environment. The majority of the studies are validated with small or single-source data, and little attention is paid to cross-dataset validation and multi-centre clinical validation. The Indian research on explainable AI is in its infancy and interpretability is usually constrained to some simple visualization, without a formalized evaluation. Furthermore, real-time implementation, computer performance, and integration into the current hospital operations are not thoroughly investigated.

Although our national research programs have shown that there is significant technical advancement in automated brain tumor analysis, it is apparent that there is a dire need of end-to-end deep learning structures that are robust, explainable, and deployable in clinical settings. The solution of these gaps will largely improve the application of AI-based brain tumor diagnostic systems into a practical medical solution to the Indian medical ecosystem.

6. INTERNATIONAL STATUS

On the international front, studies of automated brain tumor detection, segmentation, and classification with MRI have received a boost, due to the advancements in the field of deep learning, the presence of large-scale benchmark datasets, and the interdisciplinary co-operation between computer scientists and medical specialists. On a global scale, scholars have investigated

convolutional neural networks (CNNs), U-Net-based architectures, attention, vision transformers, and hybrid deep learning models in order to enhance the accuracy and robustness of the diagnostic tasks.

The use of established benchmark datasets, BraTS, TCIA, Figshare CE-MRI and institutional multi-center datasets has been integrated into the international literature, and they allow a comparative evaluation as well as methodological benchmarking. A range of best-in-use models is able to have a classification and segmentation rate of over 97 percent and advanced designs which embrace residual learning, attention gates, and transformer-based feature extractions to remember complicated tumor properties. Grad-CAM, SHAP and LIME are also explainable AI (XAI) techniques that have attracted interest in recent years to be used to improve transparency and clinician trust.

As a result of these developments, there are also unresolved issues in international research. Most of the high performing models are computationally intensive and make them impossible to implement in a real clinical environment. There is still a major concern of generalization amidst heterogeneous imaging situations, types of scanners, and patient groups. Moreover, explainability is increasingly being considered, but the absence of quantitative measures of interpretability that are standardized, and the infrequent combination of uncertainty estimation prevent clinical implementation.

There is strong methodological sophistication and performance in MRI based brain tumor analysis in international research. The necessity to have unified end to end structures that integrate segmentation and classification, robust generalization, standard explainability and real time deployment is however a fundamental research effort. Meeting such international issues is the basis and inspiring factor of the proposed PhD study.

7. STUDY HYPOTHESES

The research hypotheses of the present study are as follows and formulated based on the gaps found in the literature survey:

1) H1: The proposed Attention U-Net based on end-to-end deep learning framework with a proposed attention-based CNN based on multi-class classification, will yield higher diagnostic accuracy and consistency than the traditional multi-stage and classification-only models.
2) H2: Introduction of Explainable AI methods (Grad-CAM and LIME) into the proposed framework will not harm the classification performance but will increase the interpretability and clinical trust.

3) H3: Multi-source and heterogeneous training and validation of MRI datasets will enhance better generalization ability and strength of the proposed model under diverse imaging conditions.
4) H4: Evolutionary optimization-based means (GWO and AGWO) to be integrated into the proposed attention-based CNN will enhance convergence, decrease overfitting, and overall classification.

All these hypotheses lead to the design, implementation of the proposed research, and its evaluation.

8. STUDY RATIONALE

The MRI diagnosis of brain tumor is a very complicated and sensitive issue that requires accuracy, reliability and quick decision making. Despite the impressive results of deep learning-based solutions in automated brain tumor detection and classification, their application to practical clinical settings is still not widespread because of a number of questions that have not been answered yet. These are broken diagnostic pipelines, the inability to generalize in different imaging settings, lack of interpretability, and large computational demands. The available research concentrates on either tumor segmentation or tumor classification as a separate task, and is frequently based on multi-stage models using hand-crafted features. These methods raise the complexity of the systems and lower the diagnostic coherence of the systems. Additionally, although the recent models are claimed to be highly accurate, many are trained and tested on small datasets and their robustness and reproducibility in the real-world clinical setting is a question.

There is also another vital concern connected to the low use of explainable AI methods in a clinically significant way. Explainability techniques currently used are mostly qualitative and not extensively evaluated, not allowing clinicians to regard and apply model predictions. Moreover, the lack of estimation of uncertainty also restricts the safe use of such systems in high stakes medical decision-making. Moreover, most current state-of-the-art deep learning models have computational complexity that limits the applications in real-time or resource-intensive healthcare settings, especially in developing areas where state-of-the-art computing systems may not be easily accessible.

Hence, the justification of the study is to come up with a single, explicable, and computationally viable end-to-end framework of deep learning, which spans both brain tumor segmentation and multi-classification of MRI images. With the ability to solve the problem of interpretability, generalization, and deployability, the proposed research would bridge the gap between theoretical development and practice clinical applications, which eventually would help to make correct, reliable and trusted diagnosis of brain tumors.

9. RESEARCH METHODOLOGY

The research methodology to be proposed will adhere to the stated objectives by formulating an end-to-end, explainable, robust, and computationally efficient deep learning model to diagnose brain tumors with the help of the MRI images. The methodology is designed in clear stages in order to guarantee scientific rigour, reproducibility, and clinical relevancy.

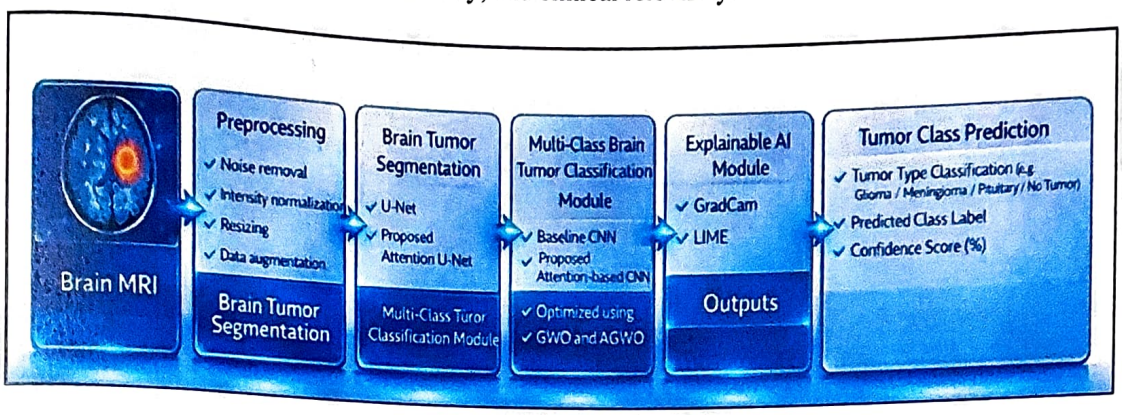


Fig.1 Block diagram of the proposed system

A. Data Collection and Dataset Preparation

Brain MRI datasets, which are publicly available, widely accepted, will be utilized, including BraTS, CE-MRI on Figshare and Kaggle brain tumor datasets. These datasets consist of multi-class tumor subsets and, in case it is provided, pixel-wise ground-truth segmentation masks. The cross-dataset validation will be done to check robustness and generalization, where datasets of various sources and scanners will be pooled.

B. Data Preprocessing

The following preprocessing steps would be done in order to have data consistency and enhance model generalization:

1. Noise Removal: Use of filtering algorithms (e.g. Gaussian/median filtering) to eliminate scanner noise and artifact.
2. Intensity Normalization: Unique pixel intensity amalgamation to diminish variability among scanners.
3. Image Resizing: Resizing all MRI images to a common resolution that can be used in deep learning.
4. Data Augmentation: To overcome the issue of class imbalance and overfitting data will be augmented with rotation, flipped, scaling and intensity variations.

C. Brain Tumor Segmentation Module

A deep learning architecture will be built with encoder-decoder that will be utilized to segment tumor at the pixel level.

1. A baseline U-Net model will be constructed and used to do comparisons.
2. An Attention U-Net architecture will be suggested to improve the performance of the tumor boundary delineation including attention mechanisms, which will pay attention to the spatial features that are relevant and eliminate the background noise.

The loss functions will be hybrid loss functions (e.g., Dice loss and cross-entropy loss) that will enhance accuracy of segmentation in case of class imbalance. The metrics of segmentation performance will be gauged on the basis of Dice coefficient, Intersection over Union (IoU), sensitivity, specificity and Hausdorff distance.

D. Multi-Class Brain Tumor Classification Module

Convolutional neural networks will be used to develop a multi-class classification framework.

1. Base CNN model will be installed in order to benchmark.
2. A better proposed attention-based CNN will be developed to attain better performance of feature discrimination and multi-class tumor classification.

In order to increase the stability of convergence and binary optimization, evolutionary optimization methods like Grey Wolf Optimization (GWO) and Adaptive Grey Wolf Optimization (AGWO) will be incorporated. Tuning of learning rates, dropout parameters, and network architecture parameters will be performed with these algorithms.

The measures of performance of classification will be accuracy, precision, recall, F1-score, AUC-ROC, and Cohen Kappa Coefficient.

E. Explainable AI (XAI) Integration

To ensure transparency and clinical interpretability, explainable AI techniques will be incorporated into the proposed framework.

1. Grad-CAM will be used to generate class-specific heatmaps highlighting important tumor regions influencing predictions.

2. LIME will be applied to provide local interpretable explanations for individual predictions.

Both qualitative visualization and quantitative interpretability assessment will be performed to evaluate the reliability and consistency of model explanations.

F. Output Generation

The final outputs of the proposed system include:

1. A **tumor segmentation mask** indicating the localized tumor region in the brain MRI
2. A **tumor class prediction**, identifying the specific brain tumor category

These outputs collectively provide comprehensive diagnostic information from a single MRI input.

10. STATISTICAL PLAN

Stage	Duration	Activity
Stage 1	Months 1–3	• Comprehensive literature survey on attention-based segmentation, optimized CNNs, evolutionary algorithms (GWO/AGWO), and XAI • Identification of research gaps in segmentation, classification, integration, and interpretability • Finalization of objectives, hypotheses, and proposed architecture • Research proposal preparation and approval
Stage 2	Months 4–6	• Dataset acquisition (BraTS, CE-MRI Figshare, Kaggle MRI datasets) • Multi-source dataset integration and annotation verification • Dataset imbalance and heterogeneity assessment • Statistical profiling of MRI distributions
Stage 3	Months 7–9	• Advanced preprocessing (noise removal, intensity normalization, resizing) • Data augmentation strategies for class imbalance handling • Preprocessing pipeline standardization across datasets

Stage 4	Months 10–12	• Development of baseline segmentation models (U-Net) • Implementation of proposed Multi-Scale Attention U-Net • Evaluation using Dice, IoU, sensitivity, specificity, Hausdorff distance
Stage 5	Months 13–15	• Development of baseline CNN classification models • Design of attention-based CNN architecture • Multi-class tumor classification evaluation (Accuracy, Precision, Recall, F1-score, AUC)
Stage 6	Months 16–18	• Integration of Grey Wolf Optimization (GWO) and Adaptive GWO (AGWO) for hyperparameter tuning • Optimization of learning rate, dropout, filters, and attention parameters
Stage 7	Months 19–21	• Development of integrated end-to-end segmentation + classification framework
Stage 8	Months 22–24	• Integration of Explainable AI (Grad-CAM, LIME) • Quantitative and qualitative interpretability assessment • Consistency evaluation of explanation maps across tumor classes
Stage 9	Months 25–27	• Cross-dataset and cross-scanner validation • Generalization and robustness evaluation • Hypothesis testing and statistical significance analysis
Stage 10	Months 28–30	• Computational optimization (model pruning, parameter reduction) • Latency, memory, and inference time analysis • Real-time feasibility evaluation for clinical deployment
Stage 11	Months 31–33	• Comparative benchmarking with state-of-the-art CNN, transformer, and hybrid models • Uncertainty estimation and confidence-aware prediction analysis
Stage 12	Months 34–36	• Final validation and result consolidation • Thesis writing and documentation • Journal and conference paper submissions • Thesis submission and completion of PhD requirements

11. LIMITATIONS OF THE STUDY

- Rich with reference to publicly available MRI data, the study cannot necessarily represent the diversity of clinical imaging scenarios, types of scanners, and the patients in the real world.
- Multi-center prospective clinical validation is outside the scope of this study, and thus, it may be possible that the proposed model is not universally applicable to other hospitals and imaging protocols.
- The methods of explainable AI (XAI) applied to the research are predominately visual and attribution methods, and the meaning of such explanations can be associated with some subjectivity because of the absence of standardized clinical interpretability measures.
- Computational efficiency Computational efficiency may be constrained, though, in real-time execution on real clinical or embedded medical hardware and may not perform as well on real-world constraints.
- The proposed model is potentially susceptible to the bias of classes and inconsistencies in the annotations of the publicly available data that cannot be eradicated despite the application of data augmentation and sound training methods.
- Certainly, the study is limited to MRI-based analysis, thus, other forms of diagnosis, including histopathology, genomic data, or clinical reports, are not included.

12. EXPECTED OUTCOMES OF THE STUDY

- Creation of a Proposed Attention-Based Segmentation Model: A powerful Attention U-Net architecture to segment brain tumors in pixels with better Dice and IoU scores in a variety of MRI images.
- Proposed Attention-Based CNN for Classification: Optimization of the attention-based CNN with high-quality multi-classification (glioma, meningioma, pituitary, and no tumor) and high-quality feature discrimination and generalization.
- Optional Governance of GWO and AGWO: A combination of Grey Wolf optimization and Adaptive GEO to provide automated hyperparameter optimization resulting in better convergence, less overfitting, and better model performance.
- The third paper is from 2019 and will focus on end-to-end, which is an integrated framework of a single deep learning system that offers segmentation and classification in a single pipeline to enhance diagnostic consistency and efficiency.

- Explainable AI Integration: Integration of Grad-CAM and LIME to get interpretable predictions, which are backed by qualitative and quantitative analysis of model explanations.

13. CONCLUSION

The proposed research will develop an optimized, end-to-end, explainable deep learning model that is robust for segmenting brain tumors and multi-classifying MRI images. This paper addresses the critical shortcoming of available methods by combining a multi-scale attention-guided encoder-decoder framework for tumor segmentation with an optimized attention-based CNN for tumor classification. The proposed framework is characterized by automated hyperparameter optimization (through the addition of Grey Wolf Optimization (GWO) and Adaptive Grey Wolf Optimization (AGWO)). The proposed framework guarantees greater stability in convergence and improved generalization.

In contrast to the fragmented pipelines that treat segmentation and classification separately, the given system provides a single segmentation-classification architecture, allowing for greater diagnostic coherence and refined feature extraction. Transparency of the model is ensured through Explainable AI approaches such as Grad-CAM and LIME, which can deliver clinically meaningful visual explanations and increase trust in automated predictions. Moreover, it has cross-dataset validation and computational optimization techniques, which guarantee robustness, scalability, and real-world clinical applicability. The suggested study adds to the optimized, interpretable, and all-encompassing AI-assisted diagnostic framework to advance reliable, clinically deployable brain tumor diagnostic systems.

14. BIBLIOGRAPHY

[1] L. A. Abraham, G. Palanisamy, G. Veerapu, and J. S. Nisha, "Exploring the potential of explainable AI in brain tumor detection and classification: A systematic review," *Artificial Intelligence Review*, vol. 59, Art. no. 14, 2026, doi: 10.1007/s10462-025-11410-8.

[2] R. Khan, S. Taj, Z. U. Khan, S. U. Khan, J. Khan, T. Arshad, and S. Ayouni, "High-precision brain tumor classification from MRI images using an advanced hybrid deep learning method with minimal radiation exposure," *Journal of Radiation Research and Applied Sciences*, vol. 18, Art. no. 101858, Aug. 2025, doi: 10.1016/j.jrras.2025.101858.

[3] I. Joy S., G. Sriram, and S. V. Sriram, "Deep CNN-based multi-grade brain tumor classification with enhanced data augmentation," *Procedia Computer Science*, vol. 260, pp. 300–307, 2025, doi: 10.1016/j.procs.2025.03.205.

[4] T. A. Fahim, F. B. Alam, and M. A. Hossain, "Brain tumor detection, classification and segmentation by deep learning models from MRI images: Recent approaches, challenges and future directions," *Array*, vol. 28, Art. no. 100571, Nov. 2025, doi: 10.1016/j.array.2025.100571.

[5] S. Kar and P. K. Singh, "MBTC-Net: Multimodal brain tumor classification from CT and MRI scans using deep neural network with multi-head attention mechanism," *Medicine in Novel Technology and Devices*, vol. 27, Art. no. 100382, 2025, doi: 10.1016/j.medntd.2025.100382.

[6] A. J. Aiya, N. Wani, M. Ramani, A. Kumar, S. Pant, K. Kotecha, A. Kulkarni, and A. Al-Danakh, "Optimized deep learning for brain tumor detection: A hybrid approach with attention mechanisms and clinical explainability," *Scientific Reports*, vol. 15, Art. no. 31386, 2025, doi: 10.1038/s41598-025-04591-3.

[7] Z. Xiong, "ResSAXU-Net for multimodal brain tumor segmentation from brain MRI," *Scientific Reports*, vol. 15, Art. no. 24179, 2025, doi: 10.1038/s41598-025-09539-1.

[8] Y. Wong, E. L. M. Su, C. F. Yeong, W. Holderbaum, and C. Yang, "Brain tumor classification using MRI images and deep learning techniques," *PLOS ONE*, vol. 20, no. 5, Art. no. e0322624, May 2025, doi: 10.1371/journal.pone.0322624.

[9] A. Ali, X. Li, W. K. Mashwani, M. Abiad, F. K. Karim, and S. M. Mostafa, "Multi-class brain tumor MRI segmentation and classification using deep learning and machine learning approaches," *Cancer Imaging*, vol. 25, Art. no. 131, 2025, doi: 10.1186/s40644-025-00953-2.

[10] A. J. Aiya, N. Wani, M. Ramani, A. Kumar, S. Pant, K. Kotecha, A. Kulkarni, and A. Al-Danakh, "Optimized deep learning for brain tumor detection: A hybrid approach with attention mechanisms and clinical explainability," *Scientific Reports*, vol. 15, Art. no. 31386, 2025, doi: 10.1038/s41598-025-04591-3.

[11] D. Ç. Ertuğrul, Ö. Toygar, and H. Karataş, "Majority voting-based tumor detection: Brain tumor detection on X-ray and MRI images using fusion of hybrid learning methods," *Acta Scientiarum. Technology*, vol. 47, Art. no. e71395, 2025, doi: 10.4025/actascitechnol.v47i1.71395.

[12] N. A. Babar, J. Lateef, S. Syed, J. Dietlmeier, N. E. O'Connor, G. B. Raupp, and A. Spanias, "Brain tumor classification in MRI scans using edge computing and a shallow attention-guided CNN," *Biomedicines*, vol. 13, no. 10, Art. no. 2571, 2025, doi: 10.3390/biomedicines13102571.

[13] D. Arif, Z. Mehmood, A. Ullah, A. Fawad, and S. Winberg, "Automated technique for brain tumor detection from magnetic resonance imaging based on local features, ensemble

classification, and YOLOv3," *BioMed Research International*, vol. 2025, Art. no. 5531209, 2025, doi: 10.1155/bmri/5531209.

[14] I. Zine-dine, J. Riffi, K. El Fazazy, I. El Batteoui, M. A. Mahraz, and H. Tairi, "A review: Machine learning techniques of brain tumor classification and segmentation," *Machine Graphics & Vision*, vol. 34, no. 3, pp. 31–55, 2025, doi: 10.22630/MGV.2025.34.3.2.

[15] S. Ghosh, S. Mazumder, P. Kundu, and A. Dey, "Developing an optimized deep learning framework for brain tumors detection and classification," *EPJ Web of Conferences*, vol. 325, Art. no. 01015, 2025, doi: 10.1051/epjconf/202532501015.

[16] A. Mjahad and A. Rosado-Muñoz, "Optimizing tumor detection in brain MRI with one-class SVM and convolutional neural network-based feature extraction," *Journal of Imaging*, vol. 11, no. 7, Art. no. 207, 2025, doi: 10.3390/jimaging11070207.

[17] M. A. Gómez-Guzmán, L. Jiménez-Beristain, E. E. García-Guerrero, O. A. Aguirre-Castro, J. J. Esqueda-Elizondo, E. R. Ramos-Acosta, G. M. Galindo-Aldana, C. Torres-Gonzalez, and E. Inzunza-Gonzalez, "Enhanced multi-class brain tumor classification in MRI using pre-trained CNNs and transformer architectures," *Technologies*, vol. 13, no. 9, Art. no. 379, Aug. 2025, doi: 10.3390/technologies13090379.

[18] M. Agarwal, G. Rani, A. Kumar, P. K. K, R. Manikandan, and A. H. Gandomi, "Deep learning for enhanced brain tumor detection and classification," *Results in Engineering*, vol. 22, Art. no. 102117, Apr. 2024, doi: 10.1016/j.rineng.2024.102117.

[19] S. Anantharajan, S. Gunasekaran, T. Subramanian, and V. R., "MRI brain tumor detection using deep learning and machine learning approaches," *Measurement: Sensors*, vol. 31, Art. no. 101026, Jan. 2024, doi: 10.1016/j.measen.2024.101026.

[20] S. G. De Benedictis, G. Gargano, and G. Settembre, "Enhanced MRI brain tumor detection and classification via topological data analysis and low-rank tensor decomposition," *Journal of Computational Mathematics and Data Science*, vol. 13, Art. no. 100103, Sep. 2024, doi: 10.1016/j.jcmds.2024.100103.

[21] L. C. S. Reddy, M. Elangovan, M. Vamsikrishna, and C. Ravindra, "Brain tumor detection and classification using deep learning models on MRI scans," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 10, Mar. 2024, doi: 10.4108/eetpht. 10.5553.

[22] M. S. Mithun and S. J. Jawhar, "Detection and classification on MRI images of brain tumor using YOLO NAS deep learning model," *Journal of Radiation Research and Applied Sciences*, vol. 17, Art. no. 101113, Oct. 2024, doi: 10.1016/j.jrras.2024.101113.

[23] M. M. Ahmed, M. M. Hossain, M. R. Islam, M. S. Ali, A. A. Nafi, M. F. Ahmed, K. M. Ahmed, M. S. Miah, M. M. Rahman, M. Niu, and M. K. Islam, "Brain tumor detection and classification in MRI using hybrid ViT and GRU model with explainable AI in Southern

Bangladesh," *Scientific Reports*, vol. 14, Art. no. 22797, 2024, doi: 10.1038/s41598-024-71893-3.

[24] T. Rahman and M. S. Islam, "MRI brain tumor detection and classification using parallel deep convolutional neural networks," *Measurement: Sensors*, vol. 26, Art. no. 100694, Feb. 2023, doi: 10.1016/j.measen.2023.100694.

[25] K. Neamah, F. Mohamed, M. M. Adnan, T. Saba, S. A. Bahaj, K. A. Kadhim, and A. R. Khan, "Brain tumor classification and detection based on deep learning models: A systematic review," *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3347545.

[26] B. B. Vimala, S. Srinivasan, S. K. Mathivanan, Mahalakshmi, P. Jayagopal, and G. T. Dalu, "Detection and classification of brain tumor using hybrid deep learning models," *Scientific Reports*, vol. 13, Art. no. 23029, 2023, doi: 10.1038/s41598-023-50505-6.